A

Data points with ISO-SPLIT hyperplane

Projection histogram with best unimodal fit and cut point

Empirical CDF with best uimodal fit

Projection direction →

Dip score = 0.6   Dip score = 0.8   Dip score = 2.7   Dip score = 4.0   Dip score = 5.3

B

C

A   B

Unit 1

Nbhd A: Keep
Nbhd B: Not detected
Nbhd C: Discard

Unit 2

Nbhd A: Discard
Nbhd B: Keep
Nbhd C: Discard

Unit 3

Nbhd A: Not detected
Nbhd B: Keep
Nbhd C: Keep

First consolidation pass

Second consolidation pass

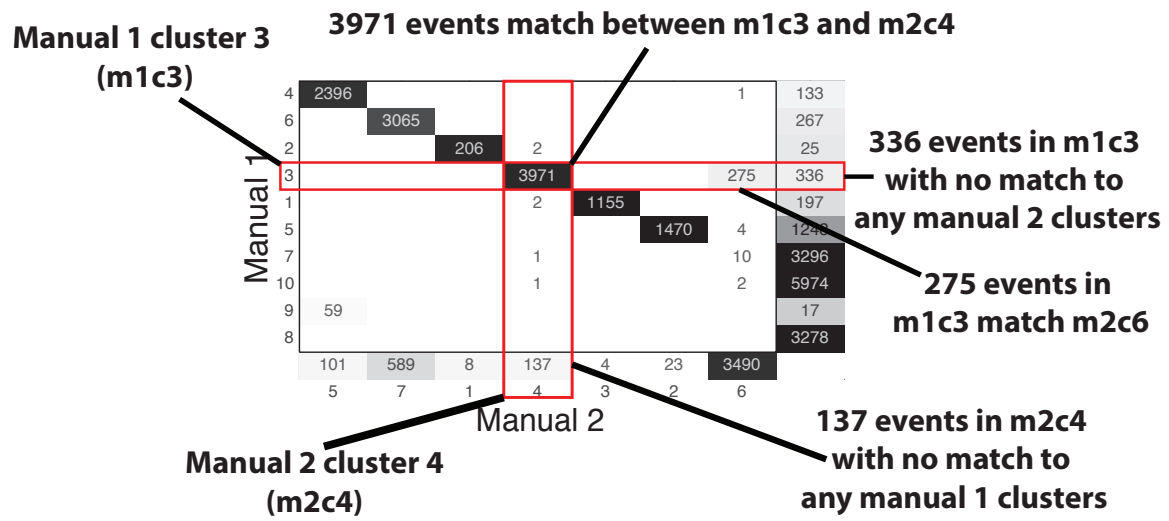Compare close match
Nbhd B: Keep
Nbhd C: Discard

Figure S1

Figure S1 (related to Figure 1). ISO-CUT and neighborhood (nbhd) consolidation illustrations.

(A) Illustration of ISO-CUT procedure which is the one-dimensional kernel operation of the ISO-SPLIT clustering algorithm. The top row shows 1200 points drawn from a mixture of two Gaussian distributions with increasing separations in the horizontal direction (separations in standard deviations are 2 through 5 in increments of 0.75). The second row from the top shows histograms of the same samples after projection onto the horizontal axis with the blue curve showing the best unimodal fit obtained using up-down isotonic regression. The corresponding empirical cumulative distributions and the best unimodal fits are shown in the third row from the top. The last three samples are considered by the test to be non-unimodal after projection, according to a dip score threshold of 1, as indicated by the red dashed lines at the cut points. (B) Illustration of the cluster consolidation stage of MountainSort where redundant clusters detected in different electrode neighborhoods are removed. The neighborhoods of electrodes A, B, and C are shown with the blue, red, and black diamonds, respectively. The illustrative units 1, 2, and 3 are detected redundantly on the channels, but only one copy of each is retained after the first and second passes of the procedure. Unit 1 is kept on channel A because its peak amplitude occurs on the channel on which it was detected. Similarly, unit 2 is kept on channel B. Unit 3 represents an ambiguous case that is handled in the second pass of the procedure.
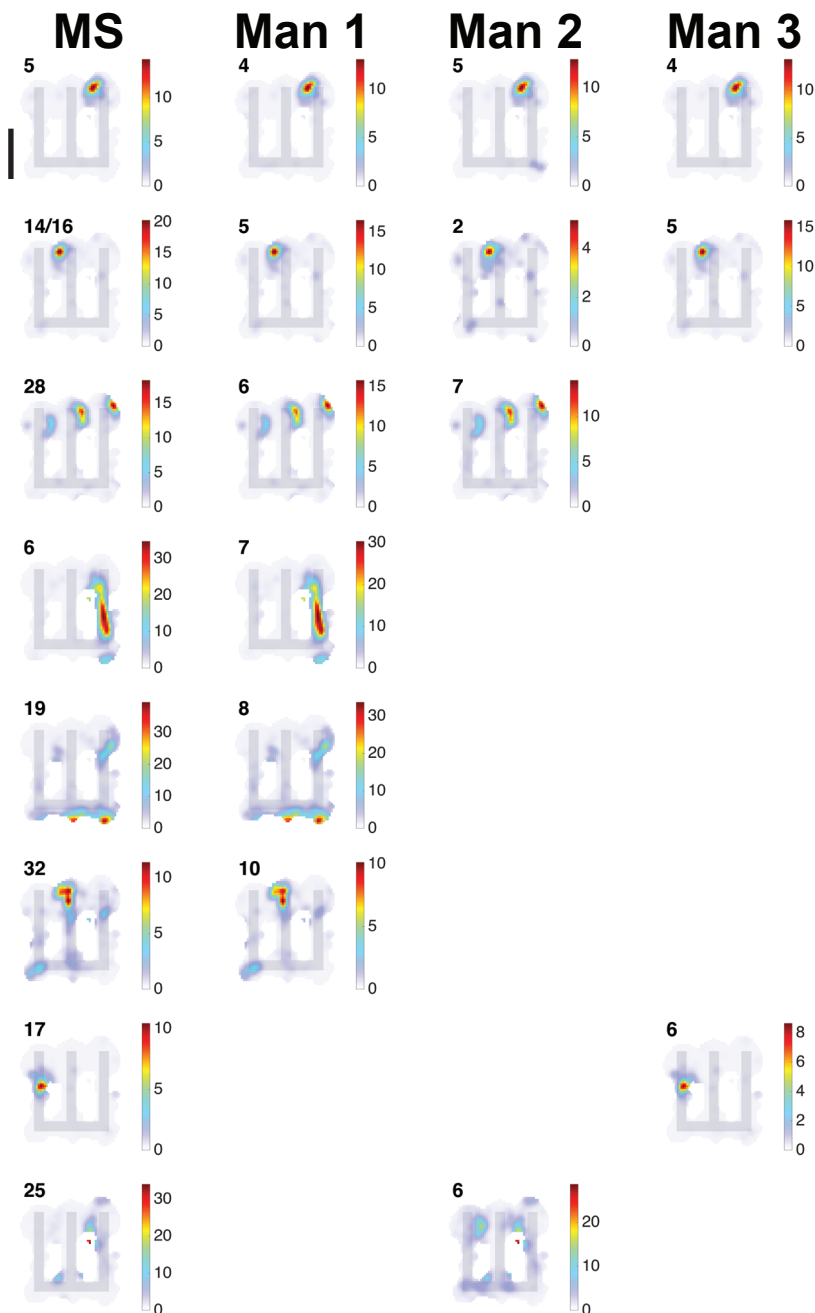
A

Manual 1 cluster 3
(m1c3)

3971 events match between m1c3 and m2c4

	5	7	1	4	3	2	6	
4	2396						1	133
6		3065						267
2			206	2				25
3				3971			275	336
1				2	1155			197
5						1470	4	124?
7				1			10	3296
10				1			2	5974
9	59							17
8								3278
	101	589	8	137	4	23	3490	

Manual 1 (rows) / **Manual 2** (columns)

336 events in m1c3
with no match to
any manual 2 clusters

275 events in
m1c3 match m2c6

Manual 2 cluster 4
(m2c4)

137 events in m2c4
with no match to
any manual 1 clusters

B

MS   Man 1   Man 2   Man 3

Figure S2

Figure S2 (related to Figure 2). Comparison among manual and MountainSort sortings.
(A) Annotated manual 1 – manual 2 confusion matrix. Red boxes highlight manual 1 cluster 3 and manual 2 cluster 4. (B) Occupancy-normalized spatial firing rate maps for all clusters found by at least one manual operator and MountainSort. Track outline is shown in gray. Scalebar corresponds to 50 cm. Also see Fig. 2C.
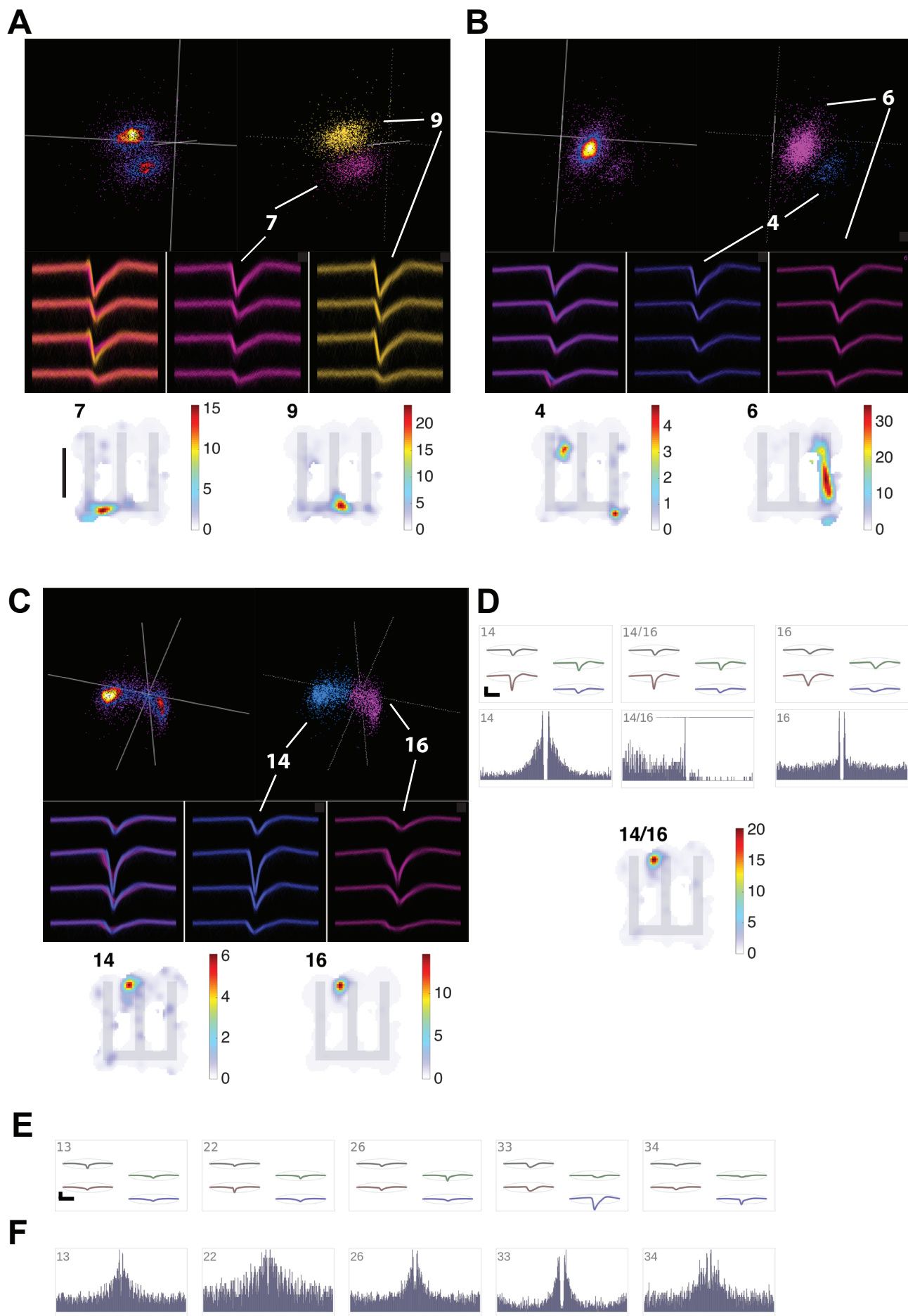
Figure S3

Figure S3 (related to Figure 3). Evaluation of cluster isolation and putative noise clusters. (A) Top, clusters shown in a rotation of the top 3 principal component dimensions for the events from MountainSort clusters 7 and 9. Top left, normalized density heat map. Top right, labeled cluster identification. Middle, an overlapping subset of 500 event waveforms from each cluster. All waveforms are bandpass filtered 300 – 6000 Hz, and have a window size of 3.33 ms. Bottom, occupancy-normalized spatial firing rate maps. Track outline is shown in gray. Scalebar corresponds to 50 cm. (B, C) As in (A), but for (B) MountainSort clusters 4 and 6, or (C) MountainSort clusters 14 and 16. (D) Top, waveforms for MountainSort clusters 14 and 16, and cluster 14/16 after automated merge. Middle, autocorrelograms for MountainSort cluster 14 (left), 16 (right), or cross correlogram (center). Bottom, occupancy-normalized spatial firing rate map for merged cluster 14/16. Track outline is shown in gray. Scalebar corresponds to 50 cm. (E) Average waveforms (bandpass filtered 300 – 6000 Hz) for the putative single-unit clusters as determined using metric thresholds: noise overlap < 0.03, isolation > 0.95, and firing rate > 0.05. Scale corresponds to 40 µv and 1 ms. (F) Autocorrelograms for the corresponding clusters; X-axis range is ±100 ms, normalized Y-axis range.
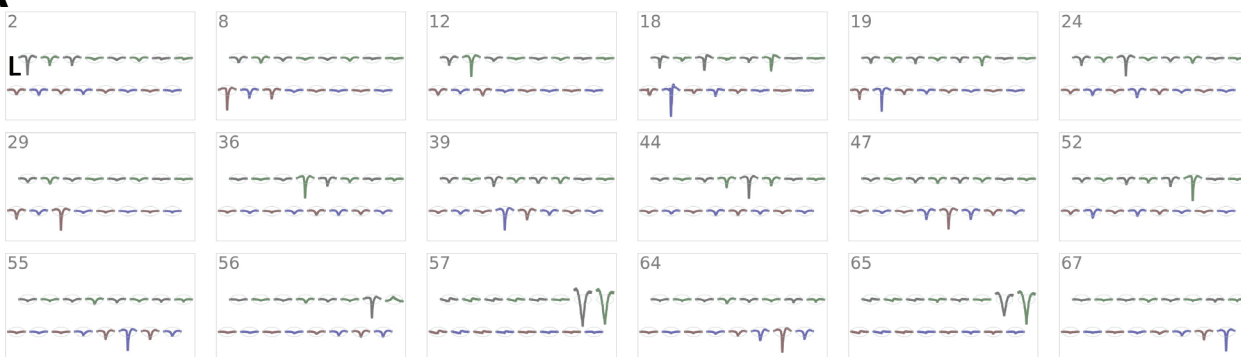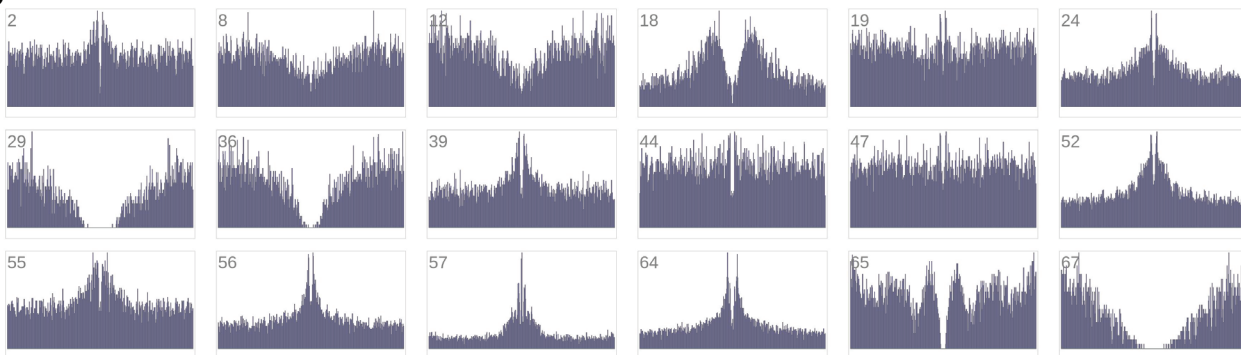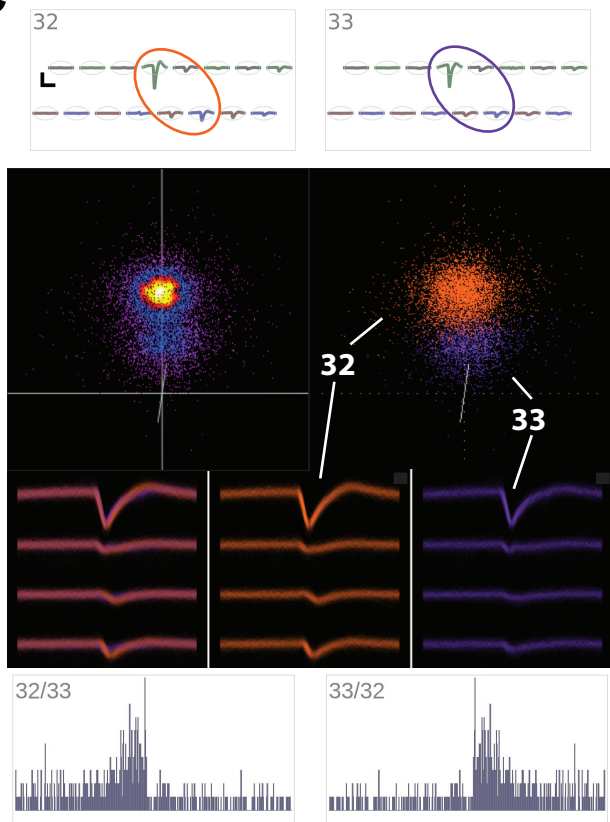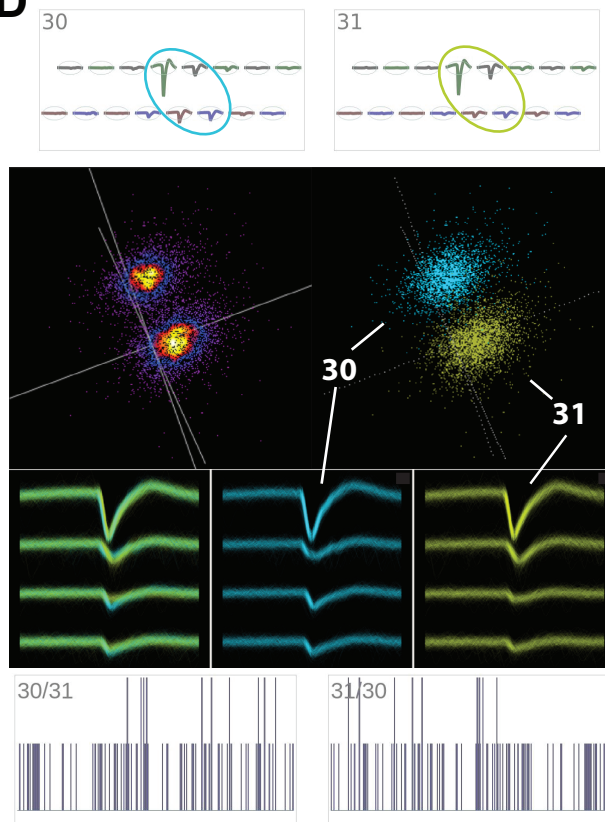
Figure S4

Figure S4 (related to Figure 4). MountainSort identification of noise, bursting, and isolated units for 16-channel probe.

(A) Units annotated as noise. Scalebar corresponds to 40 µv and 2 ms. Average waveforms filtered 300 – 6000 Hz and (B) autocorrelograms for putative noise clusters. (C) MoutainSort-identified bursting pair, clusters 32 and 33, having the lowest isolation score (0.91) of accepted clusters. Top, templates for bursting pair, scalebar corresponds to 100 µv and 2 ms. Below, clusters shown in a rotation of the top 3 principal component dimensions for the events from MountainSort clusters 32 and 33, with left, normalized density heat map and right, labeled cluster identification. Below, an overlapping subset of 500 event waveforms from each cluster. All waveforms are bandpass filtered 300 – 6000 Hz, and have a window size of 3.33 ms. Bottom, cross-correlograms with total time range of 200 ms. (D) Same as in (C), but for cluster pair 30 and 31, having second-lowest isolation score (0.97) of accepted clusters.
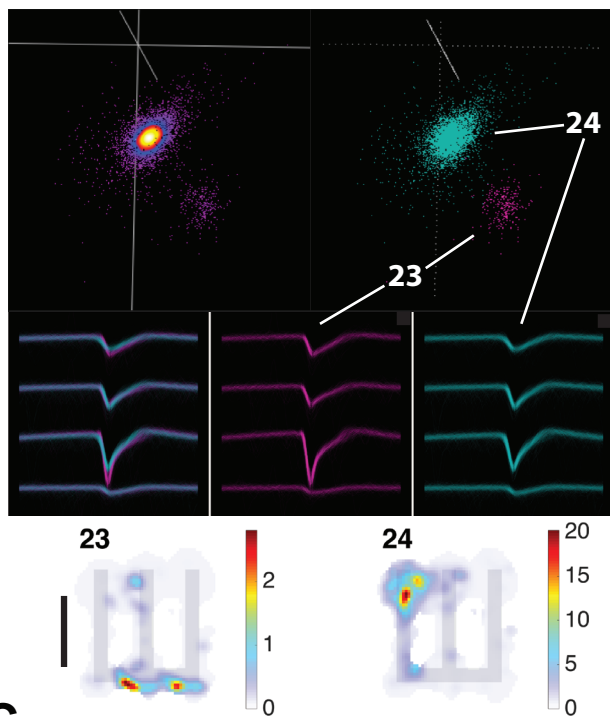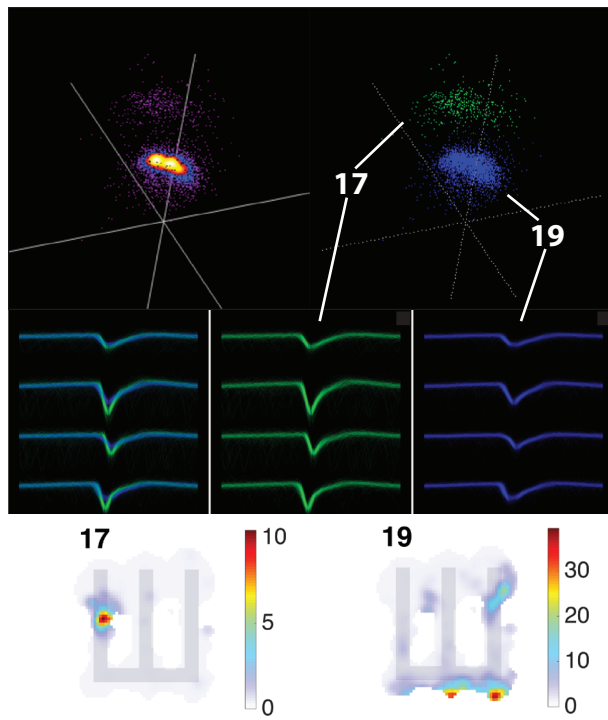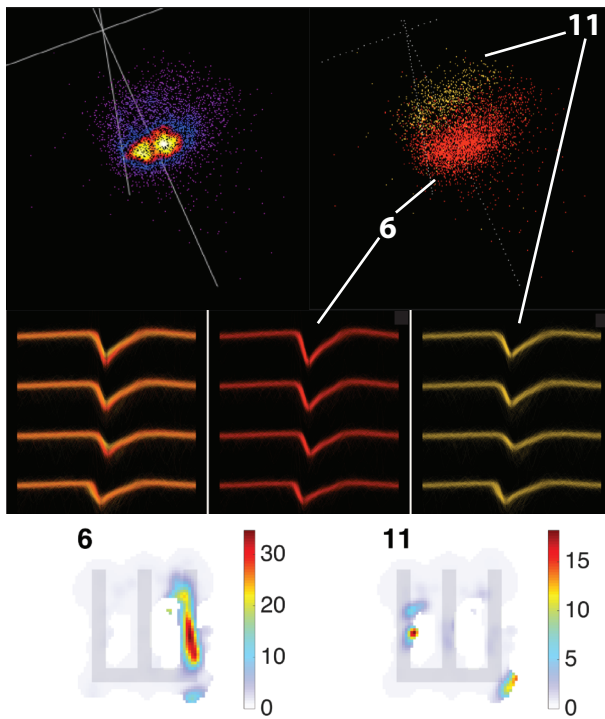
Figure S5

Figure S5 (related to Figure 5). Evaluation of cluster differences between MountainSort, Kilosort, and Spyking circus.

(A) Top, clusters shown in a rotation of the top 3 principal component dimensions for the events from MS 23 and 24, corresponding to KS 48 (failed to separate) and SC 6 (failed to separate). Top left, normalized density heat map. Top right, labeled cluster identification. Middle, an overlapping subset of 500 event waveforms from each cluster. All waveforms are bandpass filtered 300 – 6000 Hz, and have a window size of 3.33 ms. Bottom, occupancy-normalized spatial firing rate maps. Track outline is shown in gray. Scalebar corresponds to 50 cm. (B, C) As in (A), but for (B) MS 17 and 19, which form subsets of KS 8 (failed to separate) and SC 23 (failed to separate), or (C) MS 6 and 11, corresponding to KS 2 (failed to separate) and forming a subset of SC 2 (failed to separate).